Analyzing Collapsed Contingency Tables Without Actually Collapsing
Author(s): Paul D. Allison
Source: *American Sociological Review*, Vol. 45, No. 1 (Feb., 1980), pp. 123-130
Published by: American Sociological Association
Stable URL: http://www.jstor.org/stable/2095247
Accessed: 12/09/2009 20:38

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/action/showPublisher?publisherCode=asa.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit organization founded in 1995 to build trusted digital archives for scholarship. We work with the scholarly community to preserve their work and the materials they rely upon, and to build a common research platform that promotes the discovery and use of these resources. For more information about JSTOR, please contact support@jstor.org.

American Sociological Association is collaborating with JSTOR to digitize, preserve and extend access to *American Sociological Review*.

http://www.jstor.org

ANALYZING COLLAPSED CONTINGENCY TABLES WITHOUT ACTUALLY COLLAPSING*

PAUL D. ALLISON
Cornell University

American Sociological Review 1980, Vol. 45 (February):123–130

For a variety of reasons, log-linear models are sometimes fit to "collapsed" contingency tables which sum over the categories of one or more variables in the original table. I show here that the same results can be readily obtained by fitting a log-linear model to the full table, without actually collapsing it. All that is necessary is to exclude certain parameters pertaining to the collapsed variable. This approach is particularly useful for the analysis of recursive systems of categorical variables.

There are several situations in which it is desirable to take a contingency table with a given number of cells and "collapse" it into a smaller contingency table by summing over some or all of the categories of one or more variables. It may be that some of the categories are thought to be theoretically or practically equivalent. Or perhaps one of the variables is considered to be irrelevant to the analysis at hand. Most importantly, collapsing is an essential feature of Goodman's (1973) method for the log-linear analysis of recursive systems of categorical variables.

For example, suppose we have four categorical variables—A, B, C, and D—and we assume that A is causally prior to B, C, and D; B is causally prior to C and D; and C is causally prior to D. Goodman proposes that one first analyze the four-way table, with D taken as the dependent variable. Then one collapses over D to obtain a three-way table, with C taken as the dependent variable. Finally, one collapses over C to study the effect of A on B.

There are disagreements as to the appropriateness of this method. Reynolds (1977) is troubled by the fact that significant associations can appear in the collapsed table even when there are no such associations in the full table, and suggests that this may produce misleading results. However, Gillespie (1978) and Goodman (1979) have countered that in a recursive system it is the collapsed table that more accurately describes the true causal structure. Thus, when looking at the effects of A and B on C, it is misleading to further cross-classify on a subsequent dependent variable D.

Let me emphasize at the outset that I completely agree with Goodman and Gillespie. There is simply no way to represent a recursive causal system by a single log-linear model for the full multiway table. The method of successive collapsings, on the other hand, seems entirely appropriate for such recursive systems.

Nevertheless, collapsing can be tedious, especially when the original table is large. Even when one has a computer program to do the collapsing, it can be inconvenient to shift back and forth between different collapsed tables. I will show here that, while the method of collapsing is a correct way to analyze recursive systems, it is not strictly necessary. One can obtain identical results by fitting a succession of models to the full multiway table. Not only does this reduce computation and increase flexibility, but it also offers important insights into the model itself. Further, this method—while especially useful for recursive systems—also has important implications for other situations in which collapsing seems desirable. The method is based on formal results obtained by Goodman (1971) and by Bishop et al. (1975), though they did not develop these particular implications in their work. I

*Direct all communications to: Paul D. Allison; Department of Sociology; 323 Uris Hall; Cornell University; Ithaca, NY 14853.

For helpful comments, I am indebted to Roger Avery, Edward Hackett, and Reuben Snipper. An earlier version of this paper was presented at the annual meeting of the American Sociological Association, Boston, 1979.

Table 1. Attitudes toward Jobs for Blacks by Education, Region, and Year (Cell Frequencies)

		1946		1963	
		Attitude			
		Favorable	Unfavorable	Favorable	Unfavorable
South	Grade	52	206	71	44
	High	64	163	126	31
	College	42	51	55	14
Non-South	Grade	311	365	148	61
	High	439	374	410	56
	College	204	83	201	8

begin with an example but, further on, I will set down some general principles and indicate their mathematical justification.

The data to be analyzed consist of a four-way contingency table reported by Davis (1974), who adapted it from Schwartz (1967). The variables are year (1946, 1963), region (South, non-South), education (grade school, high school, college), and attitude toward jobs for blacks (favorable, unfavorable).[1] I will represent these variables by the letters A, B, C, and D, respectively, and I assume that they constitute a recursive causal system in which causal priority corresponds to alphabetical order. The table is reproduced here as Table 1.

Since the analysis of the four-way table, with D taken as dependent, is the same for both Goodman's method and the method proposed here, I do not report any of those results. Suppose, now, that we have already completed that analysis and we want to look at the effects of A and B on C. Goodman says to collapse the four-way table by summing over the categories of variable D to create a three-way table, ABC. Then, one can proceed to fit appropriate log-linear models to this collapsed table. Five models for the collapsed table are presented in Table 2A, along with their respective degrees of freedom and the likelihood ratio chi-square statistic (denoted by G^2).[2] The models are specified by the minimal set of marginal tables fitted to the data; but each of these marginal tables also corresponds to a set of parameters in the log-linear model. As is usual when one of the variables is taken to be dependent on the others, I have fitted the marginal table (AB) for the independent variables in each model. This ensures that the fitted model is equivalent to a logit model for the dependent variable. For further discussion of this point, see Fienberg (1977:79–81) or Gillespie (1978). A cursory examination of these models leads to the conclusion that model 2 is the simplest model that fits the data reasonably well. It says that A and B both have main effects on C, but there is no higher-order interaction.

Now, let us do it a different way. In Table 2B, I present results from fitting the same five models to the full four-way table. Since variable D is not included in any of the fitted marginal tables, this is equivalent to setting equal to zero all parameters pertaining to variable D in each of the fitted models. At first glance, the results seem to be markedly different. The chi-square statistics are all over 800, indicating that none of the five models fits the data well. However, a closer look reveals that *differences* in chi-square values for any two models are identical in Tables 2A and 2B, and that the same is true for differences in degrees of freedom. Moreover, since we know that the saturated model for the three-way table must have a chi-square of zero with zero degrees of freedom, we can use this as a baseline to recover the same chi-square values and degrees of freedom that we got in Table 2A. To accomplish this, we simply subtract the results for model 1 from the results for each of the succeeding models. This is shown in the last two columns of Table 2B.

[1] The last variable is the dichotomized response to the question, "Do you think Negroes should have as good a chance as white people to get any kind of job or do you think white people should have the first chance at any kind of job?"

[2] All models were fit by maximum-likelihood estimation using the program GLIM.

Table 2. Results from Fitting the Same Models to the Collapsed Table and to the Complete Table

A. Models Fitted to ABC Table				
Fitted Marginals	G^2	d.f.		
1. ABC	0	0		
2. AB, AC, BC	2.4	2		
3. AB, AC	21.4	4		
4. AB, BC	72.9	4		
5. AB, C	89.3	6		
B. Models Fitted to ABCD Table				
Fitted Marginals	G^2	d.f.	$G^2-863.0$	d.f.-12
1. ABC	863.0	12	0	0
2. AB, AC, BC	865.4	14	2.4	2
3. AB, AC	884.4	16	21.4	4
4. AB, BC	935.9	16	72.9	4
5. AB, C	952.3	18	89.3	6

These results demonstrate that we can easily fit models to collapsed tables without actually collapsing, and still get the same chi-squares. What about parameter estimates? These are also the same, even without any sort of adjustment. Table 3 gives estimated coefficients and their standardized values for model 2—the simplest model which fits the data—computed under both procedures.[3] The estimates are all identical—with the exception of the grand mean (or intercept) which has no substantive interpretation, anyway.

Table 3. Coefficient Estimates for Model 2 in Table 2*

Parameter	Estimate	Estimate/Standard Error
A	.366	4.49
B	−.997	−9.88
C1	.066	.75
C2	.826	10.65
A.B	−.216	−2.67
A.C1	.759	7.43
A.C2	.197	2.10
B.C1	.292	2.64
B.C2	−.074	−.69

* In the full table, the intercept is 4.62. In the collapsed table, the intercept is 5.31

To complete the example, let us examine the effect of variable A on variable B. Goodman says to collapse over both C and D and, then, to fit models to the two-way AB table. In Table 4, I have fitted both the saturated model (which allows for an effect of A on B) and the model of independence, which asserts that A has no effect on B. The model of independence can be rejected at the .05 level, and the estimated log-linear parameter for the effect of A on B is .17. Table 4 also displays results from fitting the same models to the full four-way table. The chi-squares are enormous, but the difference in chi-square (between the "saturated" model and the model of independence) is 4.5, the same as that in the

Table 4. Results from Fitting the Same Models to a Collapsed Table and to the Complete Table

Fitted Marginals	G^2	d.f.
Models for the AB Table:		
1. AB	0	0
2. A, B	4.5	1
Models for the ABCD Table:		
1. AB	1408.0	20
2. A, B	1412.5	21

[3] These estimates were produced by the program GLIM. Readers whose experience with programs for estimating log-linear models is limited to ECTA may find these estimates somewhat puzzling. Unlike ECTA, GLIM uses an algorithm known as iteratively reweighted least-squares which produces maximum-likelihood estimates equivalent to those produced by ECTA. However, GLIM uses a different parametrization of the model; whereas ECTA constrains the sums of certain parameters to be equal to zero, GLIM arbitrarily constrains one of the same set of parameters to be equal to zero (Holt, 1979). This difference is analogous to the difference between the parametrization of analysis of variance and the parametrization of multiple regression with a set of dummy variables as independent variables. Also, unlike ECTA, GLIM produces standard-error estimates for the parameters of unsaturated models which were used here to calculate standardized parameter values (each estimate is divided by its standard error). In spite of these differences, the point made by Table 3 is the same for GLIM or ECTA; you get the same parameter estimates whether you fit the model to the full table or to the collapsed table.

two-way table. The estimated parameter for the effect of A on B is also the same.

At this point, the general principle should be clear: to analyze a collapsed table without actually collapsing, simply fit to the full table the same marginal tables that you would have fitted to the collapsed table. The parameter estimates and their standard errors will be identical to those obtained from the collapsed table. The chi-square and the associated degrees of freedom will be higher; but differences in chi-square and degrees of freedom between models will be the same. To obtain the identical chi-square values, fit to the full table the "saturated" model for the collapsed table and then subtract its chi-square and degrees of freedom from those of the unsaturated models. Note that these results hold only for maximum likelihood estimates and for the likelihood ratio chi-square. For Pearson's goodness-of-fit chi-square, the equality will be only approximate.

Proof

At this point, the reader who is not interested in the mathematical detail may wish to skip ahead to the sections on additional applications. The equality of differences in chi-square is easily proved, using results in Goodman (1971). Consider models 1–5 in Table 2B. Goodman shows that the likelihood ratio chi-square for each of these models can be partitioned into two components: (a) the chi-square for testing the hypothesis that D is conditionally equiprobable, given variables A, B, and C; (b) the chi-square for a hypothesis about the ABC marginal table. The particular hypothesis tested in (b) corresponds to the fitted marginal tables. Since the component due to (a) is the same for all five models, it immediately follows that differences in chi-square must be equal to the differences in chi-square obtained by analyzing the ABC table directly. The generalization to higher-dimensioned tables is straightforward.

The equality of the estimated parameters is slightly more difficult to prove. Consider a three-way table for variables A, B and C. I will show that, if all parameters pertaining to variable C are set equal to zero, the estimated parameters pertaining to variables A and B are the same whether calculated from the three-way table or from the two-way table obtained by collapsing over variable C. First, some notations:

x_{ijk}—Observed cell frequency for level i of variable A, level j of B, and level k of C.

x_{ij+}—Observed cell frequency for table collapsed over C.

$\hat{m}_{ij}$—Maximum-likelihood estimate of expected cell frequency under a given model for the two-way table.

$\hat{m}_{ijk}$—Maximum-likelihood estimate of expected cell frequency under an equivalent model applied to the three-way table (all parameters pertaining to C must be set equal to 0).

$\hat{m}_{ij+}$—Frequencies obtained by collapsing $\hat{m}_{ijk}$ over C.

Estimated parameters are always algebraic functions of the estimated expected frequencies. What we have to show, therefore, is that the parameters obtained from $\hat{m}_{ijk}$ are the same as those obtained from $\hat{m}_{ij}$. The proof is in two steps. First, I show that the parameters obtained from $\hat{m}_{ijk}$ are the same as those obtained from $\hat{m}_{ij+}$. Then, I show that $\hat{m}_{ij+} = \hat{m}_{ij}$.

Step 1. In the table of frequencies, $\hat{m}_{ijk}$, C is independent of the variables A and B since we have explicitly set to zero all parameters linking C to A and B. We can therefore invoke the collapsibility theorem of Bishop et al. (1975:47), which says that if one set of variables is independent of another set, we can collapse over the first set without affecting the parameters pertaining to the second set. This means that we get the same estimated parameters from the $\hat{m}_{ij+}$ as from the $\hat{m}_{ijk}$.

Step 2. A basic property of maximum-likelihood estimation of log-linear models is that the estimated expected frequencies must sum to the fitted marginal tables (Bishop et al., 1975:69). Since we are fitting the same marginal tables to both the three-way and the two-way tables, obviously both $\hat{m}_{ijk}$ and $\hat{m}_{ij}$ must sum to the same marginal tables. It follows that $\hat{m}_{ij+}$ and $\hat{m}_{ij}$ must also sum to the same margi-

nal tables. It is also the case (Bishop et al., 1975:69) that there can be only one set of estimated cell frequencies that both satisfies the internal constraints implied by the model and sums to the fitted marginals. Hence, $\hat{m}_{ij+} = \hat{m}_{ij}$.

This result is immediately generalizable to higher dimensional tables. Suppose we have a table for variables P, Q, R, S, and T and we wish to collapse over variables S and T. We can treat S and T as a single joint variable, C, and treat P and Q as a single joint variable, A. Letting R=B, we can apply the preceding proof directly.

Collapsing over Independent Variables

To this point, I have focused on the issue of collapsing over one or more dependent variables in a recursive system. These results also shed light on what it means to collapse over an independent variable. Consider, again, the example of a four-way table with variables A, B, C, and D and suppose that D is considered to be dependent on the other three variables in a logit-type model. In such a model, it is convenient to divide the log-linear parameters into two groups:

(a) Parameters that pertain to the dependent variable and its relationship with the independent variables;
(b) Parameters that pertain only to the independent variables and the relationships among them.

Although the parameters in group (b) are always estimated in a dependent variable model, they are not ordinarily of any substantive interest (Gillespie, 1978). This is because they measure relationships that are conditional on the levels of the dependent variable and, thus, have no causal interpretation.

Now, suppose that we want to estimate models which (among other things) assert that there is no effect of C on D. There are two ways to do this:

1. Set to zero all parameters which pertain to both C and D (equivalently, do not fit any marginal tables which include both C and D as variables) for any models fitted to the ABCD table.
2. Collapse over variable C to create an ABD table; then estimate models for this new table.

These two methods are equivalent in two senses. First, the estimated effects of A and B on D will be the same for both methods. Second, tests of hypotheses about the effects of A and B on D will give identical results for both methods. In other words, any conclusions we draw about the parameters in group (a) are identical under either procedure. This will not be true of the parameters in group (b) but, as I have already indicated, these do not generally have any substantive interpretation.

To illustrate these points, let us return to the data in Table 1. In Table 5, I present the goodness-of-fit statistics for four models fitted to the ABCD table and for four models fitted to the ABD table. All of these models can be interpreted as logit models for variable D as the dependent variable. (For that to be true, it was necessary to fit the marginal table for the independent variables in each model.) Models 1–4 correspond to models 5–8, respectively, in their assertions about the effects of variables A and B on variable D. This is not immediately apparent since the chi-squares (and their associated degrees of freedom) for models 1–4 are substantially higher than those for models 5–8. Yet, the differences in chi-square (and degrees of freedom) among models 1, 2, 3, and 4 are identical to the corresponding differences among models 5, 6, 7 and 8. Thus, tests of hypotheses about the effects of A and B on D yield the same results under either procedure.

In the lower panel of Table 5, I present parameter estimates for models 2 and 6. Both models allow for main effects of variables A and B on variable D but exclude any higher-order interactions. I have separated the parameters into group (a) those which pertain to variable D, and group (b) those which do not. The group (a) estimates are clearly the same under both procedures. In group (b), corresponding parameter estimates not only differ under the two procedures, but also, there are many parameters estimated under model 2 that do not appear at all in model 6 (i.e., all the parameters pertaining to the relationships between variable C

Table 5. Results for Models which Delete the Effects of an Independent Variable

Fitted Marginals	G^2	d.f.
Models for ABCD Table:		
1. ABC, ABD	143.8	8
2. ABC, AD, BD	147.5	9
3. ABC, AD	292.6	10
4. ABC, D	738.0	11
Models for ABD Table:		
5. ABD	0	0
6. AB, AD, BD	3.7	1
7. AB, AD	148.8	2
8. AB, D	594.2	3

Parameter Estimates for Models 2 and 6:

		Model 2		Model 6	
	Parameter	Estimate	Estimate/S.E.	Estimate	Estimate/S.E.
(a)	D	1.90	22.5	1.90	22.5
	A.D	−1.78	−19.9	−1.78	−19.9
	B.D	−1.01	−11.8	−1.01	−11.8
(b)	GM	3.30	32.7	4.74	58.6
	A	1.60	13.7	1.98	23.5
	B	−.30	−1.9	−.14	−1.5
	A.B.	−.41	−2.2	−.57	−6.3
	C1	.00	0.0		
	C2	.80	9.6		
	A.C1	.86	7.1		
	A.C2	.24	2.2		
	B.C1	.51	2.8		
	B.C2	.02	.1		
	A.B.C1	−.35	−1.5		
	A.B.C2	−.17	−.8		

and variables A and D). As we have repeatedly observed, however, the group (b) parameters are not relevant to the causal model.

We have shown that models 1–4 are equivalent to models 5–8 with respect to estimating and testing the effects of A and B on D. Yet, clearly, there is an advantage to fitting models to the full table rather than to the collapsed table. While model 5 completely ignores the relationship between C and D, model 1 explicitly asserts that C has no effect on D. Hence the goodness-of-fit statistic gives a direct test of that hypothesis. With a chi-square of 143.8, and 8 degrees of freedom, this hypothesis must obviously be rejected. If we had only collapsed the table over variable C—without performing this test—we would have been led to the misleading conclusion that model 6 fits the data very well. It may fit the collapsed table, but the corresponding model does not fit the full table.

The reader may have noticed that what is being asserted here about the equivalence of collapsing over independent variables, and setting effect parameters equal to zero, differs somewhat from the earlier discussion about collapsing over dependent variables. In the earlier section, we set to zero *all* the parameters pertaining to a given variable, and we asserted that the estimates for *all* the remaining parameters would be identical to those obtained from the collapsed table. Here, however, we set to zero only those parameters describing the effect of a given variable on the dependent variable, and we assert that the parameters describing the effects of other variables on the dependent variable are the same as those obtained from the collapsed table. While this is a different claim, the proof is quite similar. The assertion that differences in chi-square are the same for models 1–4 and models 5–8 follows from Goodman's (1971) demonstration that the chi-squares in models 1–4 can be additively decomposed into two parts: (1) a chi-square for testing the hypothesis that D is conditionally independent of C, given A and B; and (2) a chi-

square for some hypothesis about the marginal table ABD. The equality of parameter estimates again follows from Bishop et al.'s (1975:47) collapsibility theorem. The proof is the same as that given earlier, except that a more general version of the collapsibility theorem is invoked.[4]

The more limited claim of this section could actually have been used in the preceding section where the aim was to collapse over dependent variables. Thus, to estimate the effects of A and B on C from the ABCD table, we could have set to zero only those parameters pertaining to both D and C instead of excluding *all* the parameters pertaining to D. The result would have been that estimates and tests of the effects of A and B on C would have been the same as those obtained from the ABC table. However, the parameters describing the relationship between A and B would have differed under the two alternative procedures.

It should also be noted that the results obtained in models 3 and 4 of Table 5 are equivalent to results that would be obtained by collapsing over both variables B and C to create a 2 × 2 table for variables A and D. Specifically, the A.D parameter estimated in model 3 is the same as that obtained from the 2 × 2 table; also, the difference in chi-square between models 3 and 4 is the same as that for the test of independence in the 2 × 2 table.

Collapsing Selected Categories

These results also apply to the case in which some, but not all, of the categories of a variable are collapsed into a single category. In Table 1, for example, education (variable C) has three categories: grade school, high school, and college. Suppose we believed that, net of other variables, there was no difference—in attitudes toward jobs for blacks—between those persons with grade-school educations and those with high-school educations. Yet, we still believed that there was a difference in attitudes between those with college educations and those without. One approach would be to collapse the grade-school and high-school categories into a single category and, then, estimate models for this collapsed table. Models 5–8 in Table 6 do just that. In this table, C′ represents the collapsed, two-category version of variable C. Alternatively, one could constrain the grade-school and high-school parameters to be equal in models for the full, uncollapsed table. While this is difficult to do with ECTA, it is easily accomplished with GLIM and certain other programs.[5] In Table 6, models 1–4 impose such constraints on the parameters for the relationship between the independent variables and the dependent variable. (No such constraints were imposed on the parameters describing the relationships among independent variables, since these have no causal interpretation.)

A comparison of models 1–4 with the corresponding models 5–8 shows that these two methods are, again, essentially equivalent. The chi-squares and degrees of freedom for models 1–4 are higher than those for the corresponding models 5–8, but differences in chi-squares and degrees of freedom within the first group are identical to the corresponding differences within the second group. Although not shown here, the estimated effects of A, B, and C′ on variable D are also the same

Table 6. Results for Models which Collapse over Selected Categories of a Single Variable

Fitted Marginals	G^2	d.f.
Models for the ABCD Table:		
1. ABC, ABC′D	52.9	4
2. ABC, ABD, AC′D, BC′D	58.6	5
3. ABC, AD, BD, C′D	64.7	8
4. ABC, AD, BD	147.6	9
Models for the ABC′D Table:		
5. ABC′D	0	0
6. ABC′, ABD, AC′D, BC′D	5.7	1
7. ABC′, AD, BD, C′D	11.8	4
8. ABC′, AD, BD	94.7	5

[4] The generalized version of the collapsibility theorem says the following: First, divide the variables into three mutually exclusive groups; if group 1 is independent of group 2 (but not necessarily group 3), we can collapse over group 1 without affecting any of the parameters pertaining to group 2. In this example, group 1 is variable C, group 2 is D, and group 3 includes A and B.

[5] Inquiries about GLIM should be addressed to Numerical Algorithms Group, Ltd.; NAG Central Office; 7 Banbury Road; Oxford OX2 6NN, U.K.

under either method. As in the preceding section, however, there is an important advantage to the method of imposing constraints on models for the full table; it provides an explicit test of the hypothesis that those constraints are true in the population. Thus, the chi-square for model 1 is a test statistic for the hypothesis that there is no difference—in attitudes about jobs for blacks—between those persons with grade-school and those with high-school educations. Clearly, that hypothesis can be rejected. Models 5–8, on the other hand, have nothing at all to say about that hypothesis.

REFERENCES

Bishop, Yvonne M. M., Stephen E. Fienberg, and Paul W. Holland
1975 Discrete Multivariate Analysis. Cambridge, Mass.: MIT Press.

Davis, James A.
1974 "Hierarchical models for significance tests in multivariate contingency tables." Pp. 189–231 in Herbert Costner (ed.), Sociological Methodology 1973–74. San Francisco: Jossey-Bass.

Fienberg, Stephen E.
1977 The Analysis of Cross-Classified Categorical Data. Cambridge, Mass.: MIT Press.

Gillepsie, Michael W.
1978 "The application of log-linear techniques to recursive models: comment on Reynolds." American Journal of Sociology 84:718–22.

Goodman, Leo A.
1971 "Partitioning of chi-square, analysis of marginal contingency tables, and estimation of expected frequencies in multidimensional contingency tables." Journal of the American Statistical Association 66:339–92.
1973 "The analysis of multidimensional tables when some variables are posterior to others: a modified path analysis approach." Biometrika 60:179–92.
1979 "A brief guide to the causal analysis of data from surveys." American Journal of Sociology 84:1078–96.

Holt, D.
1979 "Log-linear models for contingency table analysis: on the interpretation of parameters." Sociological Methods and Research 7:330–6.

Reynolds, H. T.
1977 "Some comments on the causal analysis of surveys with log-linear models." American Journal of Sociology 83:127–43.

Schwartz, M. A.
1967 Trends in White Attitudes toward Negroes. University of Chicago, NORC Report No. 119.